

Proceso y metodología de validación de las Estadísticas de Transporte -ESTRA

Octubre, 2022

Contenido

Proceso y metodología de validación de las Estadísticas de Transporte – ESTRA3	
1. Introducción	3
2. Objetivo	3
3. Etapas de validación	3
4. Recolección.....	3
4.1 Crítica e integración de la base de datos.....	4
4.2 Clasificaciones y/o codificaciones	5
Proceso y metodología de las distintas etapas de validación de Vehículos Matriculados	7
1. Ordenar la BDD e identificar casos duplicados.....	7
2. Validación de la variable clase de vehículo	10
3. Arreglo carga y pasajeros	14
Imputación.....	15
4. CAPACIDAD_PASAJEROS - Capacidad de pasajeros.....	15
5. CAPACIDAD_DE_CARGA Capacidad de carga	15
6. Imputación en STATA	16
Proceso y metodología de las distintas etapas de validación de Sinistros de Tránsito.....	18
1. Creación de variables (SPSS)	18
Proceso y metodología de las distintas etapas de validación de Transporte Ferroviario.....	22
1. Cambios y creación de variables	22
Proceso y metodología de las distintas etapas de validación de Transporte Aéreo	24
.....	25
Proceso y metodología de las distintas etapas de validación de Transporte Marítimo.....	26
1. Creación y validación de variables.....	26
5. Comparación de la información	29
6. Conclusiones.....	31

Proceso y metodología de validación de las Estadísticas de Transporte – ESTRA

1. Introducción

En vista de la necesidad de información básica, actualizada y completa que le permita al gobierno de Ecuador planear una adecuada política de desarrollo y organizar adecuadamente el sistema de transportación a nivel nacional; el Anuario de Estadísticas de Transporte - ANET, se ha realizado desde el año de 1966 al año de 1974 y de 1976 hasta la actualidad; su primera edición se publicó bajo el nombre Estadística de Transporte 1965 – 1966.

El ANET tiene como objetivo presentar información a partir de registros administrativos (RA) de diferentes instituciones, a fin de ofrecer una amplia descripción del sector del transporte en el país en ámbitos como: número de vehículos matriculados, pasajeros transportados por ferrocarril, carga y pasajeros transportados por vía aérea y marítima, así como, siniestros de tránsito a nivel nacional cuyas fuentes de información son las siguientes entidades gubernamentales:

- La Agencia Nacional de Tránsito - ANT
- Empresa Ferrocarriles del Ecuador Empresa Pública - FEEP
- Dirección General de Aviación Civil – DGAC
- Subsecretaría de Puertos y Transporte Marítimo y Fluvial – SPTMF

En este contexto, existe la necesidad de definir efectivamente los pasos para el procesamiento de información de los diferentes componentes del anuario para la optimización de tiempos y solucionar posibles problemas que se presenten. Esto mediante un proceso de validación y verificación de la información; y obtener bases de datos consistentes para la generación de resultados.

2. Objetivo

Diseñar un documento de procesos y establecer la metodología de las distintas etapas de validación a las cuales debe ser sometida la información recibida para la construcción del ANET.

3. Etapas de validación

Las etapas de validación y verificación a las cuales debe ser sometida la información recibida, se encuentran detallados en la fase de recolección, procesamiento y análisis de los datos de a las actividades del cronograma. Una vez que la información es recibida se realiza la planificación de la validación de acuerdo al orden de llegada de la data de los diferentes componentes.

4. Recolección

La actividad previa al procesamiento estadístico es la transferencia y recolección de datos desde las fuentes de registros administrativos. Para esto, se establecen escenarios de transferencia (identificación de herramientas para entrega - recepción de información y actores que intervienen) que permiten el acceso oportuno a los datos y se preserve los principios de confidencialidad, integridad y disponibilidad de la información.

Tabla 1. Calendarización de solicitud y entrega de información por fuentes

Fuentes de información		Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic
Vehículos Matriculados	Solicita												
	Recibe												
Transporte Aéreo	Solicita												
	Recibe												
Transporte Marítimo y fluvial	Solicita												
	Recibe												
Siniestros de tránsito	Solicita												
	Recibe												

Fuente: Instituto Nacional de Estadística y Censos (INEC)

4.1 Crítica e integración de la base de datos

Para la crítica de ESTRA 2021 se considera las siguientes acciones:

- Revisión de formatos establecidos para la entrega de la información.
- Revisión del número de variables requeridas.
- Revisión de datos omitidos.

La integración de las bases de datos se realiza de manera homologada en función de la información entregada por cada una de las fuentes de información:

Tabla 2. Proceso de integración de bases de datos - ESTRA

Fuentes de Información	Proceso de integración de bases de datos ¹
Vehículos Matriculados	Es la integración de las bases de datos entregadas en formato XLS y CSV para luego exportar en SPSS.
Transporte Marítimo y fluvial	Es la integración de tabulados entregados en formato XLS para luego exportar en SPSS.
Siniestros de tránsito	Es la integración de bases de datos entregados en formato XLS y CSV para luego consolidar la información y exportar en SPSS.

Fuente: Instituto Nacional de Estadística y Censos (INEC)

Cabe señalar que la información presentada de Transporte Aéreo no dispone de bases de datos como tal, sino que se construye en base al Boletín Estadístico de Tráfico Aéreo de la DGAC.

¹ Las estadísticas de vía aérea provienen del boletín estadístico que publica la DGAC más no entregan bases de datos.

4.2 Clasificaciones y/o codificaciones

Para la clasificación y codificación de las variables suministradas en los distintos registros administrativos, se realiza y/o actualiza las sintaxis de codificación (en SPSS, R o STATA), a fin de garantizar homologación y consistencia (con las publicaciones de años anteriores) en la asignación de las categorías requeridas para la presentación de resultados, por ejemplo: clase de vehículos, tipo de cargas, otras.

En esta etapa se utiliza el **Clasificador Geográfico Estadístico 2021**, especialmente para las variables de tipo geográfico como: provincia y cantón.

A continuación, se describen las actividades realizadas para la generación y revisión de las bases de datos por componente. Cabe señalar que este documento debe ser complementado con la sintaxis ya que tiene mayor detalle.

01.

Vehículos Matriculados

Proceso y metodología de las distintas etapas de validación de Vehículos Matriculados

1. Ordenar la BDD e identificar casos duplicados

Ordenar los casos por placa de forma ascendente.

PLACA	IDENTIFICACION	NOMBRE_PROPIETARIO
A001669	702344565	GONZALEZ VASQUEZ PROI
A001794	105280309	BARRERA SUAREZ MAURIC
A002382	1707335244	ESPINOZA HEREDIA JOSE
A003086	917920019	GARCIA VERA JOSE VITER
A004070	301156261	ZHANGALLIMBAY CHICAIZA
A004111	105813661	PADILLA BALAREZO LUIS A
A004370	925091514	JARA PESANTES CRISTHIA
A004474	927344499	SALGADO CAMEJO JEFFEI

SORT CASES BY PLACA (A).

EXECUTE.

Una vez ordenada la base de datos se procede a identificar casos duplicados en la base mediante la aplicación de sintaxis; en la cual se genera el reporte.

1 Duplicado

0 Registros únicos.

Estadísticos		
testd Registros duplicados		
N	Válido	2361175
	Perdidos	0

- En caso de presentar casos de duplicados se procede a verificar en la base, y si la base contiene la misma información en todas las variables con el mismo identificador, una vez verificado, se realiza la depuración de los datos.
- Se suele solicitar una base semestral y adicionalmente una base anual, se comparan y unen las bases para tomar casos que pudieron omitirse y posteriormente se obtiene la base para iniciar la validación, depuración y procesamiento.

Para la bdd 2021, existieron casos duplicados y omitidos, una vez identificados y una vez unificada la base, se considera apta para iniciar con el proceso de validación, imputación, etc.

- Se genera la variable MES, tomado de la *Fecha de transacción de la base*.

```
STRING MES (A10).
COMPUTE MES=CHAR.SUBSTR(FECHA_DE_TRANSACCION,4,2).
VARIABLE LABELS MES 'Mes matriculación'.
EXECUTE.

RECODE MES ("Ene"=01) ("Feb"=02) ("Mar"=03) ("Abr"=04) ("May"=05) ("Jun"=06) ("Jul"=07) ("Ago"=08) ("Sep"=09) ("Oct"=10)
("Nov"=11) ("Dic"=12) INTO MES.
VARIABLE LABELS MES 'Mes Matr'.
FORMATS MES (F2.0).
VALUE LABELS MES
```

- Se crea la variable **aniofab** basándose en la variable original **año**. Algunos casos no tienen la información sobre el año de fabricación del vehículo. Para

corregir esos casos y completar faltantes, se debe utilizar la matrícula e investigar en el sistema del SRI para saber el año correcto, así como la página de la ANT:

- ✓ <https://srienlinea.sri.gob.ec/sri-en-linea/SriVehiculosWeb/ConsultaValoresPagarVehiculo/Consultas/consultaRubros>

En la página se accede a información por Placa, RAMV o CPN:

- ✓ https://consultaweb.ant.gob.ec/PortalWEB/paginas/clientes/clp_criterio_consulta.jsp

En la página de consulta de multas se puede acceder a los datos del vehículo por RUC, pasaporte, cedula de identidad o placa.

- Crear la variable **prov_dom** recodificando la variable original usando el codificador geográfico. Para datos missing colocamos 99 SIN REGISTRO.

- Crear la variable **canton_dom** recodificando la variable original usando el codificador geográfico. Para datos missing colocamos 99 SIN REGISTRO.
 - Crear la variable **prov_maticu** recodificando la variable original usando el codificador geográfico. Para datos missing colocamos 99 SIN REGISTRO.
 - Crear la variable **canton_maticu** recodificando la variable original usando el codificador geográfico. Para datos missing colocamos 99 SIN REGISTRO.
- ✓ Para la creación y codificación de esas variables se trabaja con el Clasificador Geográfico Estadístico – DPA

DPA	PROV	CANTON	DPA_DESCR	DPA_FAMROG	DPA_DESCR	DPA_FAMROG	DPA_DESCR	DPA_FAMROG	DPA_DESCR	DPA_FAMROG
01	AZUAY	CUENCA	0101	CUENCA	0101	CUENCA	0101	CUENCA	0101	CUENCA
02	AZUAY	CUENCA	0102	CUENCA	0102	CUENCA	0102	CUENCA	0102	CUENCA
03	AZUAY	CUENCA	0103	CUENCA	0103	CUENCA	0103	CUENCA	0103	CUENCA
04	AZUAY	CUENCA	0104	CUENCA	0104	CUENCA	0104	CUENCA	0104	CUENCA
05	AZUAY	CUENCA	0105	CUENCA	0105	CUENCA	0105	CUENCA	0105	CUENCA
06	AZUAY	CUENCA	0106	CUENCA	0106	CUENCA	0106	CUENCA	0106	CUENCA
07	AZUAY	CUENCA	0107	CUENCA	0107	CUENCA	0107	CUENCA	0107	CUENCA
08	AZUAY	CUENCA	0108	CUENCA	0108	CUENCA	0108	CUENCA	0108	CUENCA
09	AZUAY	CUENCA	0109	CUENCA	0109	CUENCA	0109	CUENCA	0109	CUENCA
10	AZUAY	CUENCA	0110	CUENCA	0110	CUENCA	0110	CUENCA	0110	CUENCA
11	AZUAY	CUENCA	0111	CUENCA	0111	CUENCA	0111	CUENCA	0111	CUENCA
12	AZUAY	CUENCA	0112	CUENCA	0112	CUENCA	0112	CUENCA	0112	CUENCA
13	AZUAY	CUENCA	0113	CUENCA	0113	CUENCA	0113	CUENCA	0113	CUENCA
14	AZUAY	CUENCA	0114	CUENCA	0114	CUENCA	0114	CUENCA	0114	CUENCA
15	AZUAY	CUENCA	0115	CUENCA	0115	CUENCA	0115	CUENCA	0115	CUENCA
16	AZUAY	CUENCA	0116	CUENCA	0116	CUENCA	0116	CUENCA	0116	CUENCA
17	AZUAY	CUENCA	0117	CUENCA	0117	CUENCA	0117	CUENCA	0117	CUENCA
18	AZUAY	CUENCA	0118	CUENCA	0118	CUENCA	0118	CUENCA	0118	CUENCA
19	AZUAY	CUENCA	0119	CUENCA	0119	CUENCA	0119	CUENCA	0119	CUENCA
20	AZUAY	CUENCA	0120	CUENCA	0120	CUENCA	0120	CUENCA	0120	CUENCA
21	AZUAY	CUENCA	0121	CUENCA	0121	CUENCA	0121	CUENCA	0121	CUENCA
22	AZUAY	CUENCA	0122	CUENCA	0122	CUENCA	0122	CUENCA	0122	CUENCA
23	AZUAY	CUENCA	0123	CUENCA	0123	CUENCA	0123	CUENCA	0123	CUENCA
24	AZUAY	CUENCA	0124	CUENCA	0124	CUENCA	0124	CUENCA	0124	CUENCA
25	AZUAY	CUENCA	0125	CUENCA	0125	CUENCA	0125	CUENCA	0125	CUENCA
26	AZUAY	CUENCA	0126	CUENCA	0126	CUENCA	0126	CUENCA	0126	CUENCA
27	AZUAY	CUENCA	0127	CUENCA	0127	CUENCA	0127	CUENCA	0127	CUENCA
28	AZUAY	CUENCA	0128	CUENCA	0128	CUENCA	0128	CUENCA	0128	CUENCA
29	AZUAY	CUENCA	0129	CUENCA	0129	CUENCA	0129	CUENCA	0129	CUENCA
30	AZUAY	CUENCA	0130	CUENCA	0130	CUENCA	0130	CUENCA	0130	CUENCA
31	AZUAY	CUENCA	0131	CUENCA	0131	CUENCA	0131	CUENCA	0131	CUENCA
32	AZUAY	CUENCA	0132	CUENCA	0132	CUENCA	0132	CUENCA	0132	CUENCA
33	AZUAY	CUENCA	0133	CUENCA	0133	CUENCA	0133	CUENCA	0133	CUENCA
34	AZUAY	CUENCA	0134	CUENCA	0134	CUENCA	0134	CUENCA	0134	CUENCA
35	AZUAY	CUENCA	0135	CUENCA	0135	CUENCA	0135	CUENCA	0135	CUENCA
36	AZUAY	CUENCA	0136	CUENCA	0136	CUENCA	0136	CUENCA	0136	CUENCA
37	AZUAY	CUENCA	0137	CUENCA	0137	CUENCA	0137	CUENCA	0137	CUENCA
38	AZUAY	CUENCA	0138	CUENCA	0138	CUENCA	0138	CUENCA	0138	CUENCA
39	AZUAY	CUENCA	0139	CUENCA	0139	CUENCA	0139	CUENCA	0139	CUENCA
40	AZUAY	CUENCA	0140	CUENCA	0140	CUENCA	0140	CUENCA	0140	CUENCA
41	AZUAY	CUENCA	0141	CUENCA	0141	CUENCA	0141	CUENCA	0141	CUENCA
42	AZUAY	CUENCA	0142	CUENCA	0142	CUENCA	0142	CUENCA	0142	CUENCA
43	AZUAY	CUENCA	0143	CUENCA	0143	CUENCA	0143	CUENCA	0143	CUENCA
44	AZUAY	CUENCA	0144	CUENCA	0144	CUENCA	0144	CUENCA	0144	CUENCA
45	AZUAY	CUENCA	0145	CUENCA	0145	CUENCA	0145	CUENCA	0145	CUENCA
46	AZUAY	CUENCA	0146	CUENCA	0146	CUENCA	0146	CUENCA	0146	CUENCA
47	AZUAY	CUENCA	0147	CUENCA	0147	CUENCA	0147	CUENCA	0147	CUENCA
48	AZUAY	CUENCA	0148	CUENCA	0148	CUENCA	0148	CUENCA	0148	CUENCA
49	AZUAY	CUENCA	0149	CUENCA	0149	CUENCA	0149	CUENCA	0149	CUENCA
50	AZUAY	CUENCA	0150	CUENCA	0150	CUENCA	0150	CUENCA	0150	CUENCA
51	AZUAY	CUENCA	0151	CUENCA	0151	CUENCA	0151	CUENCA	0151	CUENCA
52	AZUAY	CUENCA	0152	CUENCA	0152	CUENCA	0152	CUENCA	0152	CUENCA
53	AZUAY	CUENCA	0153	CUENCA	0153	CUENCA	0153	CUENCA	0153	CUENCA
54	AZUAY	CUENCA	0154	CUENCA	0154	CUENCA	0154	CUENCA	0154	CUENCA
55	AZUAY	CUENCA	0155	CUENCA	0155	CUENCA	0155	CUENCA	0155	CUENCA
56	AZUAY	CUENCA	0156	CUENCA	0156	CUENCA	0156	CUENCA	0156	CUENCA
57	AZUAY	CUENCA	0157	CUENCA	0157	CUENCA	0157	CUENCA	0157	CUENCA
58	AZUAY	CUENCA	0158	CUENCA	0158	CUENCA	0158	CUENCA	0158	CUENCA
59	AZUAY	CUENCA	0159	CUENCA	0159	CUENCA	0159	CUENCA	0159	CUENCA
60	AZUAY	CUENCA	0160	CUENCA	0160	CUENCA	0160	CUENCA	0160	CUENCA
61	AZUAY	CUENCA	0161	CUENCA	0161	CUENCA	0161	CUENCA	0161	CUENCA
62	AZUAY	CUENCA	0162	CUENCA	0162	CUENCA	0162	CUENCA	0162	CUENCA
63	AZUAY	CUENCA	0163	CUENCA	0163	CUENCA	0163	CUENCA	0163	CUENCA
64	AZUAY	CUENCA	0164	CUENCA	0164	CUENCA	0164	CUENCA	0164	CUENCA
65	AZUAY	CUENCA	0165	CUENCA	0165	CUENCA	0165	CUENCA	0165	CUENCA
66	AZUAY	CUENCA	0166	CUENCA	0166	CUENCA	0166	CUENCA	0166	CUENCA
67	AZUAY	CUENCA	0167	CUENCA	0167	CUENCA	0167	CUENCA	0167	CUENCA
68	AZUAY	CUENCA	0168	CUENCA	0168	CUENCA	0168	CUENCA	0168	CUENCA
69	AZUAY	CUENCA	0169	CUENCA	0169	CUENCA	0169	CUENCA	0169	CUENCA
70	AZUAY	CUENCA	0170	CUENCA	0170	CUENCA	0170	CUENCA	0170	CUENCA
71	AZUAY	CUENCA	0171	CUENCA	0171	CUENCA	0171	CUENCA	0171	CUENCA
72	AZUAY	CUENCA	0172	CUENCA	0172	CUENCA	0172	CUENCA	0172	CUENCA
73	AZUAY	CUENCA	0173	CUENCA	0173	CUENCA	0173	CUENCA	0173	CUENCA
74	AZUAY	CUENCA	0174	CUENCA	0174	CUENCA	0174	CUENCA	0174	CUENCA
75	AZUAY	CUENCA	0175	CUENCA	0175	CUENCA	0175	CUENCA	0175	CUENCA
76	AZUAY	CUENCA	0176	CUENCA	0176	CUENCA	0176	CUENCA	0176	CUENCA
77	AZUAY	CUENCA	0177	CUENCA	0177	CUENCA	0177	CUENCA	0177	CUENCA
78	AZUAY	CUENCA	0178	CUENCA	0178	CUENCA	0178	CUENCA	0178	CUENCA
79	AZUAY	CUENCA	0179	CUENCA	0179	CUENCA	0179	CUENCA	0179	CUENCA
80	AZUAY	CUENCA	0180	CUENCA	0180	CUENCA	0180	CUENCA	0180	CUENCA
81	AZUAY	CUENCA	0181	CUENCA	0181	CUENCA	0181	CUENCA	0181	CUENCA
82	AZUAY	CUENCA	0182	CUENCA	0182	CUENCA	0182	CUENCA	0182	CUENCA
83	AZUAY	CUENCA	0183	CUENCA	0183	CUENCA	0183	CUENCA	0183	CUENCA
84	AZUAY	CUENCA	0184	CUENCA	0184	CUENCA	0184	CUENCA	0184	CUENCA
85	AZUAY	CUENCA	0185	CUENCA	0185	CUENCA	0185	CUENCA	0185	CUENCA
86	AZUAY	CUENCA	0186	CUENCA	0186	CUENCA	0186	CUENCA	0186	CUENCA
87	AZUAY	CUENCA	0187	CUENCA	0187	CUENCA	0187	CUENCA	0187	CUENCA
88	AZUAY	CUENCA	0188	CUENCA	0188	CUENCA	0188	CUENCA	0188	CUENCA
89	AZUAY	CUENCA	0189	CUENCA	0189	CUENCA	0189	CUENCA	0189	CUENCA
90	AZUAY	CUENCA	0190	CUENCA	0190	CUENCA	0190	CUENCA	0190	CUENCA
91	AZUAY	CUENCA	0191	CUENCA	0191	CUENCA	0191	CUENCA	0191	CUENCA
92	AZUAY	CUENCA	0192	CUENCA	0192	CUENCA	0192	CUENCA	0192	CUENCA
93	AZUAY	CUENCA	0193	CUENCA	0193	CUENCA	0193	CUENCA	0193	CUENCA
94	AZUAY	CUENCA	0194	CUENCA	0194	CUENCA	0194	CUENCA	0194	CUENCA
95	AZUAY	CUENCA	0195	CUENCA	0195	CUENCA	0195	CUENCA	0195	CUENCA
96	AZUAY	CUENCA	0196	CUENCA	0196	CUENCA	0196	CUENCA	0196	CUENCA
97	AZUAY	CUENCA	0197	CUENCA	0197	CUENCA	0197	CUENCA	0197	CUENCA
98	AZUAY	CUENCA	0198	CUENCA	0198	CUENCA	0198	CUENCA	0198	CUENCA
99	AZUAY	CUENCA	0199	CUENCA	0199	CUENCA	0199	CUENCA	0199	CUENCA

La codificación utilizada por el INEC, se encuentra estandarizada con el Registro Civil y demás entidades del Sector Público que manejan un sistema con las siguientes características:

- Se asignan los dos primeros códigos de la izquierda al código de provincia.
- El código 01 es asignado al cantón, cuya cabecera es también capital provincial.
- Dentro de cada cantón se codifican las parroquias urbanas en orden alfabético desde 01 hasta 49, asignando el 50 a la cabecera cantonal.
- A las parroquias rurales se les ordena alfabéticamente y se asigna el código respectivo del 51 al 99.

Las provincias permanecen codificadas en orden alfabético, a excepción de Galápagos, Sucumbíos, Orellana, Santo Domingo de los Tsáchilas y Santa Elena, creadas en fechas posteriores al establecimiento de este criterio de codificación; a éstas se les asigna el código inmediato superior al que tiene la última provincia, de acuerdo a la fecha de su respectiva creación.

Nota 2021: Existen 14 datos (placa) sin registro, que en la primera base entregada están correctos, se procede a tomar esos datos para completar la información de la data.

Además, se completan con datos de bases de años anteriores y datos de SRI y ANT.

Para facilitar la escritura de sintaxis, se recomienda hacer uso de un archivo Excel y trabajar con la función **concatenar**, para conseguir ahorrar tiempo.

	Recuento				
CAMILO PONCE ENRÍQUEZ	4173	CAMIL	01	0115	F (prov_domicilio=01 & CANTON_DOMICILIO='CAMILO PONCE ENRÍQUEZ') cant_domicilio=01
CHORDELEG	1554	CHORD	01	0111	F (prov_domicilio=01 & CANTON_DOMICILIO='CHORDELEG') cant_domicilio=01
CUENCA	132772	CUENC	01	0101	F (prov_domicilio=01 & CANTON_DOMICILIO='CUENCA') cant_domicilio=01
EL PAN	469	EL PA	01	0112	F (prov_domicilio=01 & CANTON_DOMICILIO='EL PAN') cant_domicilio=01
GIRÓN	2549	GIRÓN	01	0102	F (prov_domicilio=01 & CANTON_DOMICILIO='GIRÓN') cant_domicilio=01
GUACHAPALA	505	GUACH	01	0114	F (prov_domicilio=01 & CANTON_DOMICILIO='GUACHAPALA') cant_domicilio=01
GUALACEO	6855	GUALA	01	0103	F (prov_domicilio=01 & CANTON_DOMICILIO='GUALACEO') cant_domicilio=01
NABÓN	1282	NABÓN	01	0104	F (prov_domicilio=01 & CANTON_DOMICILIO='NABÓN') cant_domicilio=01
OÑA	341	OÑA	01	0110	F (prov_domicilio=01 & CANTON_DOMICILIO='OÑA') cant_domicilio=01

- Crear la variable **servicio** recodificando la variable original. Se consolidan en las categorías:
 - 1 ESTADO
 - 2 USO PÚBLICO, ALQUILER O COMERCIAL
 - 3 PARTICULAR
 - 4 OTROS.
- Para los valores missings, se investiga el modelo del vehículo en las páginas mencionadas y se asigna un uso. Para casos, donde no se pueda realizar por ese medio, se toma la información de bases de años anteriores si se dispone.
- Crear la variable **tipo_combustible** recodificando los datos de la variable. De tal manera que se dispondrá de las siguientes categorías:
 - 1 DIESEL
 - 2 GASOLINA
 - 3 HÍBRIDO
 - 4 ELÉCTRICO
 - 5 GAS LICUADO DE PETROLEO
 - 6 OTRO
 - 99 NO REPORTADO.

2. Validación de la variable clase de vehículo

Para la creación de la clase de vehículo, partir de las variables validadas en años anteriores se trabaja con la placa, se toma la clase final de 2020, luego de 2019, luego la de 2018 teniendo así: clase2020_3, clase2019_3 y case 2018_3.

NOTA: Considerar, desde este punto, trabajar con una base exclusiva para la variable clase, que incluya variables de validación como placa, marca, modelo, identificación, vin, motor, color y carga y pasajeros.

Todo esto se trabaja con las sintaxis en SPSS donde también existen notas.

- Crear la variable clase_2021 partiendo de la clase original de la base, esto permite tener la clase original. Teniendo estas variables, procedemos a identificar los casos que deberían entrar a revisión y ser considerados para la reclasificación. Un proceso que se realiza manualmente.

- Crear variable clase2021_2, tomando los valores de la variable clase2021_3, luego de la clase2020_3, luego los casos que queden vacíos se toman los valores de la clase2019_3, luego clase2018_3.
- De esa manera, existirán casos vacíos para los que se deberán tomar los valores originales, en este caso año 2021. Esto ayuda a disminuir los casos sobre los cuales se debe trabajar y categorizar.

Así habrán quedado pocos casos para trabajar en ellos y asignar categorías.

Seguido, asignamos a la variable clase:

- 1 "AUTOMÓVIL"
- 2 "AUTOBÚS"
- 3 "CAMIÓN"
- 4 "CAMIONETA"
- 5 "FURGONETA"
- 6 "SUV"
- 7 "MOTOCICLETA"
- 8 "TANQUERO"
- 9 "TRAILER"
- 10 "VOLQUETA"
- 11 "OTRA CLASE"
- 99 "NO REGISTRADO".

Al momento de categorizar adecuadamente los vehículos, en lugar de **DO IF**, que toma a todos los vehículos y los categoría automáticamente, prefírase crear un filtro para verificar que, por ejemplo, en automóvil que se encuentre dentro de la: clase2021_3=1 & pasajero>10 esté categorizándose de forma correcta.

Si suele suceder que uno o dos casos están mal categorizados, asignados de marca errónea, y por ejemplo, el automóvil en realidad era una moto o camioneta. Lo mismo para buses y otras clases de vehículos.

NOTA: Si las clases de todos los años tomados coinciden la categoría está correcta, sin embargo sucede que por x motivo la categorización pudo ser errónea. Por lo que si ninguno de los años coincide, ese caso también debe ser tomado en cuenta para revisión.

En ese caso, iniciamos con la correcta escritura de Marcas

MARCA:

AMP	2	OTRAS MARCAS	IF (MARCA='AMP') marca2021='OTRAS MARCAS'.
AMP-AMERICA MAX POWER	9	OTRAS MARCAS	IF (MARCA='AMP-AMERICA MAX POWER') marca2021='OTRAS MARCAS'.
ANDINO	1	OTRAS MARCAS	IF (MARCA='ANDINO') marca2021='OTRAS MARCAS'.
ANGLIA	4	OTRAS MARCAS	IF (MARCA='ANGLIA') marca2021='OTRAS MARCAS'.
ANKAI	13	OTRAS MARCAS	IF (MARCA='ANKAI') marca2021='OTRAS MARCAS'.
ANWA	1	OTRAS MARCAS	IF (MARCA='ANWA') marca2021='OTRAS MARCAS'.
APOLLO	9	OTRAS MARCAS	IF (MARCA='APOLLO') marca2021='OTRAS MARCAS'.
APOLLO ORION RX 250	7	OTRAS MARCAS	IF (MARCA='APOLLO ORION RX 250') marca2021='OTRAS MARCAS'.
APRILIA	333	APRILIA	IF (MARCA='APRILIA') marca2021='APRILIA'.
ARN	2	OTRAS MARCAS	IF (MARCA='ARN') marca2021='OTRAS MARCAS'.
ARO	6	OTRAS MARCAS	IF (MARCA='ARO') marca2021='OTRAS MARCAS'.
ARO CROSS LANDER	1	OTRAS MARCAS	IF (MARCA='ARO CROSS LANDER') marca2021='OTRAS MARCAS'.
ASIA	278	ASIA	IF (MARCA='ASIA') marca2021='ASIA'.
ASPT	1	OTRAS MARCAS	IF (MARCA='ASPT') marca2021='OTRAS MARCAS'.
ATUL	677	ATUL	IF (MARCA='ATUL') marca2021='ATUL'.
ATV	1	OTRAS MARCAS	IF (MARCA='ATV') marca2021='OTRAS MARCAS'.
AUDI	1906	AUDI	IF (MARCA='AUDI') marca2021='AUDI'.

El siguiente paso es generar segmentos de los casos en los que se trabajarán para realizar validaciones por caso o casos, dependiendo de la marca y el modelo.

Identificar los casos para asignar clase y abrir un documento auxiliar de sintaxis en Excel.

PLACA	marca2020	TIPO VEHICULO	MODELO_COMPLETO	clase2020_3	pasajeros Capacidad de pasajeros	tonelaje tonelaje	nueva clase	sintaxis
AAU0240	MERCEDES BENZ	CAMION	OF 1721/52	5	40,00	7,00		2 IF (PLACA="AAU0240" AND marca2020="marca2020" AND clase2020_3=3) clase2020_3=3
CAE0007	HINO	CAMION	GH1JMSA	5	42,00	12,00		2 IF (PLACA="CAE0007" AND marca2020="MERCEDES BENZ" AND clase2020_3=3) clase2020_3=3
CAE0140	CHEVROLET	BUS	FTR32M CHASIS TORPEDO	5	40,00	8,00		2 IF (PLACA="CAE0140" AND marca2020="HINO" AND clase2020_3=3) clase2020_3=3
EAA2321	HINO	BUS	AK&JRSA 7.7 3P 4X2 TM DIESEL	5	40,00	7,00		2 IF (PLACA="EAA2321" AND marca2020="CHEVROLET" AND clase2020_3=3) clase2020_3=3
EAG0574	CHEVROLET	BUS	FTR32M CHASIS TORPEDO	5	40,00	6,00		2 IF (PLACA="EAG0574" AND marca2020="HINO" AND clase2020_3=3) clase2020_3=3

Considerar validar los casos usando: PLACA, MARCA, TIPO DE VEHICULO, MODELO CLASE2021_3, PASAJEROS Y CARGA. Al momento de asignar el nuevo valor de clase, estas variables ayudan a determinar de mejor manera el vehículo.

Para sintaxis, preferir realizar el cambio desde PLACA y aplicando filtros sobre estos, ya que usando marca modelo y clase, tienden a realizar cambios sobre los vehículos no considerados para el cambio y que están asignados correctamente.

Cada vez que se validen la clase del vehículo y se corra la sintaxis de cambios, se debe generar varias veces el tabulado 8 que es marca de vehículos por clase, este sirve para identificar los casos donde se debe poner especial cuidado, porque existen marcas que no fabrican cierto tipo de vehículos.

	1 AUTOMÓ VIL	2 AUTOBÚS TA	3 CAMIÓN TA	4 CAMIONE TA	5 FURGONE TA	6 SUV	7 MOTOCIC LETA	8 TANQUE RO	9 TRAILER	10 VOLQUE TA	11 OTRA CLAS
Total	748.373	22.679	101.495	428.248	47.145	436.197	722.581	2.324	12.449	10.084	6.442
CHEVROLET	331.102	2.070	31.184	153.628	8.099	107.748	-	347	321	980	834
SUZUKI	22.601	-	-	88	302	44.439	101.717	-	-	-	16
HYUNDAI	79.993	1.042	6.399	689	13.532	51.356	-	61	57	118	355
KIA	81.548	1	3.178	1.233	7.156	50.535	-	2	-	3	64
TOYOTA	29.730	44	1.291	74.080	1.732	34.670	-	8	-	10	211
MAZDA	17.814	-	137	81.219	400	6.955	-	-	-	1	50
NISSAN	41.815	25	802	30.274	819	24.450	-	51	33	476	179
SHINERAY	-	-	-	1.390	719	1.525	74.058	-	-	-	2

Sin embargo, hay casos de vehículos muy antiguos que cierta marca dejo de producir pero que en el país siguen circulando.

También existen marca que están iniciando en la fabricación, de por ejemplo, motocicletas y antes no lo hacía. O de SUV cuando antes sólo tenían camiones. En esos casos, las páginas oficiales de las marcas son de gran ayuda para validar que el vehículo está correcto.

La validación va de la mano con la constante generación de este tabulado y cuando se tengan estos resultados se pasa a la imputación de datos que se basa en la clase del vehículo, la cual se explica más adelante.

	1 AUTOMÓ VIL	2 AUTOBÚS TA	3 CAMIÓN TA	4 CAMIONE TA	5 FURGONE TA	6 SUV	7 MOTOCIC LETA	8 TANQUE RO	9 TRAILER	10 VOLQUE TA	11 OTRA CLAS
Total	748.373	22.679	101.495	428.248	47.145	436.197	722.581	2.324	12.449	10.084	6.442
CHEVROLET	331.102	2.070	31.184	153.628	8.099	107.748	-	347	321	980	834
SUZUKI	22.601	-	-	88	302	44.439	101.717	-	-	-	16
HYUNDAI	79.993	1.042	6.399	689	13.532	51.356	-	61	57	118	355
KIA	81.548	1	3.178	1.233	7.156	50.535	-	2	-	3	64
TOYOTA	29.730	44	1.291	74.080	1.732	34.670	-	8	-	10	211
MAZDA	17.814	-	137	81.219	400	6.955	-	-	-	1	50
NISSAN	41.815	25	802	30.274	819	24.450	-	51	33	476	179
SHINERAY	-	-	-	1.390	719	1.525	74.058	-	-	-	2
FORD	5.509	80	1.963	30.009	314	21.240	-	176	115	477	524
HINO	-	12.089	32.197	-	-	-	941	1.150	3.849	1.182	-
DATONIA	-	-	-	-	-	46.567	-	-	-	-	-
BAJAJ	-	-	-	-	-	45.878	-	-	-	-	-
RANGER	-	-	-	-	-	45.309	-	-	-	-	-
HONDA	1.781	-	22	224	3.782	38.997	-	-	-	-	17
VOLKSWAGEN	32.115	2.141	237	4.750	736	2.813	42.774	-	33	51	137
MOTOR UNO	-	-	-	-	-	-	-	-	-	-	-
RENAULT	30.821	-	6	8	23	8.721	-	4	24	77	29
TAMARCA	3.974	-	3	15.645	-	16.639	-	-	-	-	217
GREAT WALL	4.542	329	5.945	8.347	444	11.643	-	49	6	107	92
NISSAN	-	-	-	-	-	30.742	-	-	-	-	-

Ahora, como ejemplo y como recomendación, a continuación se muestra cómo realizar el arreglo de las motos. Es preferible iniciar por este segmento ya que, generalmente, existen muchas marcas que sólo producen esta clase de vehículo.

- **Genere el filtro** para generar una tabla sólo de marcas de motos y expórtela a Excel:

Placa	Marca	Modelo del vehículo	Clase de vehículo original s sin arreglo	clase2021_3	Nueva clase 2021_3	
HA313B	SUKIDA	GNE 125	OTRA CLASE	11	7	IF (MARCA='SUKIDA' AND clase2021_3=11) clase2021_3=7.
HBD9303	KENBO	S3L LUXURY AC 1.5 5P 4X2 TM	SUV	6	ok	
HBD9310	KENBO	S3L LUXURY AC 1.5 5P 4X2 TM	SUV	6	ok	
HF880H	RANGER	200-6	SUV	6	7	IF (PLACA='HF880H' AND clase2021_3=6) clase2021_3=7.
HT939U	RANGER	200GY-8	MOTOCICLETA	6	7	IF (PLACA='HT939U' AND clase2021_3=6) clase2021_3=7.

Como se observa, la marca *RANGER*, puede prestar a confusiones ya que se parece a una modelo de vehículo llamado *ranger*. Es así que, motocicletas fueron colocadas como SUV.

Nos valemos del archivo Excel para facilitar la creación de sintaxis.

Nota: Se recomienda generar tabulados únicamente para motocicletas y facilitar la identificación de casos y creación de sintaxis:

	1 AUTOMÓVIL	2 AUTOBÚS	3 CAMIÓN	4 CAMIONETA	5 FURGONETA	6 SUV	7 MOTOCICLETA	8 TANQUERO	9 TRAILER	10 VOLQUETA	11 OTRA CLASE
BAJAJ	-	-	-	-	-	-	45.877	-	-	-	1
DAYTONA	1	-	-	-	-	1	46.564	-	-	-	1
DUKARE	-	-	-	-	-	-	21.138	-	-	-	1
KENBO	-	-	-	-	-	23	326	-	-	-	-
KINGDOM	1	-	-	-	-	-	386	-	-	-	-
LONGIN	-	-	1	-	-	-	17.369	-	-	-	-
MOTOR UNO	1	-	-	-	-	-	42.773	-	-	-	-
NANFANG	-	-	-	-	-	-	88	-	-	1	-
NIMBUS	-	-	-	1	-	-	606	-	-	-	-
RANGER	-	-	-	1	-	8	45.301	-	-	-	-
SUKIDA	2	-	-	-	-	-	18.152	-	-	-	1
SYM	-	-	-	-	-	1	106	-	-	-	-
TRAXX	-	-	-	1	-	-	5.300	-	1	-	-
TUNDRA	-	-	1	-	-	-	30.740	-	-	-	1
VYCAST	-	-	-	1	-	-	1.585	-	-	-	-
YAMAHA	-	-	-	-	-	1	39.571	-	-	-	-
Z1	1	-	-	-	-	1	4.560	-	-	-	-

- Si en una marca se identifica pocas inconsistencias, se puede arreglar directamente filtrando la marca y la clase del vehículo que presenta la novedad, como se muestra el siguiente caso:

IF (marca2021='FOTON' and MODELO='BJ1039V3JD3-1' and clase2021_3=3) clase2021_3=4.
 IF (marca2021='FOTON' and MODELO='BJ1028V3AA2-1' and clase2021_3=3) clase2021_3=4.
 IF (marca2021='FOTON' and MODELO='BJ1043V9JE6-1' and clase2021_3=3) clase2021_3=4.
 IF (marca2021='FOTON' and MODELO='AUMARK E BJ1044 AC 2.8 2P 4X2 TM DIESEL CN' and clase2021_3=3) clase2021_3=4.
 IF (marca2021='FOTON' and MODELO='BJ1028V3AA2-1' and clase2021_3=3) clase2021_3=4.
 IF (marca2021='FOTON' and MODELO='BJ1028V3PA2' and clase2021_3=3) clase2021_3=4.
 IF (marca2021='FOTON' and MODELO='BJ1039V3JD3-1' and clase2021_3=3) clase2021_3=4.
 IF (marca2021='FOTON' and MODELO='BJ1028V3PA2' and clase2021_3=3) clase2021_3=4.
 IF (marca2021='FOTON' and MODELO='BJ1043V9JE6-1' and clase2021_3=3) clase2021_3=4.
 IF (marca2021='FOTON' and MODELO='VIEW CS2 WIDE BJ6549 AC 2.8 4P 4X2 TM DIESEL' and clase2021_3=3) clase2021_3=5.
 IF (marca2021='FOTON' and MODELO='BJ6536B1DDA-3' and clase2021_3=2) clase2021_3=5.
 IF (marca2021='FOTON' and MODELO='VIEW C BJ6536 AC 2.8 4P 4X2 TM DIESEL' and clase2021_3=3) clase2021_3=5.
 IF (marca2021='FOTON' and MODELO='AUMAN BJ1166 AC 6.7 2P 4X2 TM DIESEL' and clase2021_3=3) clase2021_3=8.
 IF (marca2021='FOTON' and MODELO='AUMAN BJ4259 AC 11.8 2P 6X4 TM DIESEL' and clase2021_3=3) clase2021_3=8.

- En el camino de la asignación correcta de clase, se deberá identificar las marcas a cambiar. Se recomienda realizar el momento de la identificación,

ya que puede prestar a confusiones más adelante. Como se muestra en la siguiente imagen, se considera la PLACA para realizar el cambio.

```
*****CAMBIO DE MARCAS 2021*****
IF (PLACA='GZY0626') marca2021='SYM'.
IF (PLACA='HD223K') marca2021='SYM'.
IF (PLACA='IH695V') marca2021='THUNDER'.
IF (PLACA='IE482A') marca2021='Z1'.
IF (PLACA='GZY0626') marca2021='SYM'.
```

3. Arreglo carga y pasajeros

Este paso de la validación, se enfoca a la corrección de carga y pasajeros. En esta etapa se identifican inconsistencias de la carga y pasajeros con la clase del vehículo, por lo que será necesario continuar con los cambios de la manera en que se trabajó en el paso anterior.

Se recomienda usar la NORMA TÉCNICA ECUATORIANA - NTE INEN 2656 de CLASIFICACIÓN VEHICULAR. Además de los listados de homologación vehicular de la ANT que se puedan encontrar ya que contienen la clase y tipo de vehículo de acuerdo a lo establecido por esa entidad. Esto también facilita la clasificación correcta de los vehículos.

- Iniciamos con la corrección de la carga mediante sintaxis para todos los vehículos detallados, ejemplo:

```
*****AUTOMOVIL*****
```

```
do if (clase2021_3=1 & tonelajefn=75).
compute tonelajefn=0.75.
end if.
execute.
```

```
*****SUV*****
```

```
do if (clase2021_3=6 & tonelajefn=75).
compute tonelajefn=0.75.
end if.
execute.

do if (clase2021_3=6 & tonelajefn=99).
compute tonelajefn=0.99.
end if.
execute.
```

Esto permite corregir errores de digitación que pueden venir en la base.

- **Pasajeros:** Revisar, filtrar y construir sintaxis de arreglo por clase de vehículo. Por ejemplo, iniciamos con los vehículos de clase AUTOMOVIL, que tienen una capacidad de más de 10.

```
***ATOMOVIL***
```

```
USE ALL.
COMPUTE filter_$=(clase2021_3=1 & pasajero>10).
VARIABLE LABELS filter_$ 'clase2021_3=1 & pasajero>10 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.

CTABLES
/VLABELS VARIABLES=filter_$ PLACA marca2021 TIPO_VEHICULO MODELO clase2021_3 pasajero
tonelajefn
DISPLAY=BOTH
/TABLE filter_$ > PLACA > marca2021 > TIPO_VEHICULO > MODELO > clase2021_3 > pasajero
[C] > tonelajefn [COUNT F40.0]
/CATEGORIES VARIABLES=filter_$ PLACA marca2021 TIPO_VEHICULO MODELO clase2021_3
pasajero tonelajefn ORDER=A KEY=VALUE EMPTY=EXCLUDE
/CRITERIA CILEVEL=95.
```

La sintaxis, filtra los vehículos que cumplen esa característica y genera un tabulado para revisión de los casos en Excel y construir la sintaxis de arreglo:

Placa AUTOMOVILES	Tipo de vehículo	Marca	Modelo del vehículo	clase2021_3	Capacidad de pasajeros	tonelajeTn	Nueva clase	Sintaxis	Sintaxis 2
AH0094	FURGONETA	NISSAN	VANETTE	1	11,00	1,00	5	IF (PLACA=AH0094' and clase2021_3	IF (MARCA=NISSAN' and MODELO=VA
ACM0458	FURGONETA	MAZDA	E2000	1	15,00	1,00	5	IF (PLACA=ACM0458' and clase2021_3	IF (MARCA=MAZDA' and MODELO=E2
GIG0958	SEDAN	FORD	AEROSTAR WAGON XLT	1	12,00	1,00	5	IF (PLACA=GIG0958' and clase2021_3	IF (MARCA=FORD' and MODELO=AER
GLT0598	SEDAN	HYUNDAI	MATRIX STD	1	25,00	,75	5	IF (PLACA=GLT0598' and clase2021_3	IF (MARCA=HYUNDAI' and MODELO=1
GMH0070	SEDAN	CHEVROLET	20238	1	56,00	,75	pasajero 5	IF (PLACA=GMH0070' and clase2021_3	IF (MARCA=CHEVROLET' and MODEL
GMZ0809	SEDAN	CHEVROLET	CORSA EVOLUTION 4P 1.8 AC	1	75,00	1,00	pasajero 5	IF (PLACA=GMZ0809' and clase2021_3	IF (MARCA=CHEVROLET' and MODEL
GSB4030	SEDAN	KIA	RIO STYLUS LS AC	1	61733,00	,75	pasajero 5	IF (PLACA=GSB4030' and clase2021_3	IF (MARCA=KIA' and MODELO=RIO ST
GSC2253	SEDAN	CHEVROLET	AVEO ACTIVO 1.6L 4P AC	1	2011,00	,50	pasajero 5	IF (PLACA=GSC2253' and clase2021_3	IF (MARCA=CHEVROLET' and MODEL
GSC6072	SEDAN	CHEVROLET	OPTRA ADVANCE 1.8L 4P TM	1	12,00	,75	pasajero 5	IF (PLACA=GSC6072' and clase2021_3	IF (MARCA=CHEVROLET' and MODEL

- Corra la sintaxis y continúe con la siguiente categoría establecida en la sintaxis.

Luego de haber realizado este procedimiento para todas las clases de vehículos, será necesario verificar una vez más si se corrigieron las inconsistencias de clasificación. Luego, se deberá trabajar en cada caso de forma particular. De lo contrario ya se puede imputar la capacidad de pasajeros y de carga.

Se sugiere que, si necesita corregir la capacidad de carga o pasajeros lo realice de una vez en la sintaxis, esto ayudará a que corra mejor la imputación.

Imputación

4. CAPACIDAD_PASAJEROS - Capacidad de pasajeros.

Crear la variable **pasajeroT** partiendo de la variable original.

Antes de reemplazar caracteres extras como comas o espacios, se sugiere retirar mediante la función reemplazar en Excel una vez que generamos en SPSS una tabla de frecuencia de pasajeros y la llevamos a Excel para determinar los cambios que se deben realizar con sintaxis, por ejemplo de letras y espacios, doble comas y otros, ya que reemplazar directamente con sintaxis crea muchos errores que afectan la imputación.

Por ejemplo: 4-5p puede generar una capacidad de pasajeros de 45, cuando en realidad era de entre 4 a 5 pasajeros.

Posterior a estos reemplazos se pueden retirar de forma global los otros errores de digitación.

5. CAPACIDAD_DE_CARGA Capacidad de carga

Crear la variable **PESOT** partiendo de la variable original.

Antes de reemplazar caracteres extra como comas o espacios, se sugiere retirar mediante la función reemplazar en Excel, letras y espacios, doble comas y otros, ya que reemplazar directamente con sintaxis crea muchos errores que afectan la imputación.

Por ejemplo: 5 0 puede generar un peso de 50 una vez retirado el espacio, lo que en realidad corresponde a 5,0.

CAPACIDAD_PASAJEROS_COMPLETO	CAPACIDAD_DE_CARGA
5,00	,5
5,00	0,375P
5,00	,75
5,00	,75

Posterior a estos reemplazos se pueden retirar de forma global los otros errores de digitación.

Para llevar a cabo la imputación, se debe realizar mediante STATA con la sintaxis ya establecida del método de imputación que es una variante del modelo Hot-Deck. Este método de imputación, básicamente llena los espacios faltantes duplicando los valores preexistentes en la base.

En la imputación puede ocurrir que no se corra la sintaxis sobre algunos vehículos, si eso sucede, se debe identificar los casos y corregir el peso, pasajeros o clase de vehículo con sintaxis para que la imputación no arroje errores.

6. Imputación en STATA

La imputación en STATA es una sintaxis construida por Franklin Tenesaca basándose en el modelo Hot Deck. Lo que se realiza es correr la sintaxis una vez se tiene trabajada la variable clase.

- La imputación puede ser realizada varias veces ya que suele arrojar errores y requiere varias revisiones.
- Una vez imputados los datos en STATA, se traslada a formato SPSS para ser incorporados en la base donde se trabajó la clase.
- Unificar estas variables a la base original y generar la base preliminar para correr sintaxis y generar los tabulados.
- Este proceso garantiza que tanto la base trabajada como la base original tengan las variables de micro dato con las variables trabajadas finales.

02.

Siniestros de Tránsito

Proceso y metodología de las distintas etapas de validación de Sinistros de Tránsito.

Previo al traslado de la base de Excel a SPSS, se debe realizar algunos trabajos en la información en ese estado. Como por ejemplo, verificar que los totales y casos sean correctos y que coincidan con lo publicado en la página de ANT.



Es importante señalar que se debe replicar los tabulados de la ANT y verificar coincidencia de sumas y totales.

F12																	
=BUSCAR(Y12:ZONA PLANIF(SB52:SC52322;FALSO))																	
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
MES	AÑO	CÓDIGO	PAÍS DE ORIGEN	PROVENIENCIA	ZONA DE ORIGEN	INA	FECHA	NOVA	PERIODO	PERIODO	CÓDIGO CAUSA	CATEGORÍA PROBLEMA	CATEGORÍA	CAUSA DEL PROBLEMA	CLASE PRINC.	CLASE SEC	CLASE DEL PROBLEMA
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	01/01/20	01	01	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	05/01/20	05	05	C23	NO SE PUEDEN LLEVAR	4	NEUTRE LA PISTA	CHOCOS	1	CHOCOS
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	05/01/20	05	05	C23	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	04/01/20	04	04	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	04/01/20	04	04	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	07/01/20	07	07	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	07/01/20	07	07	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	04/01/20	04	04	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	04/01/20	04	04	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/20	03/01/20	03	03	CM	CONCEJO DE SATEIN	4	IMPEDIR EL EMPLEO	ATROPELLON	1	ATROPELLON
ENERO	2020	TE	ESPANOLA	BLAVIAS	INA	matrícula	08/01/2										

Con las validaciones realizadas se crea el archivo que se enviará a ANT sí existen diferencias en algún dato para ser aclarado y/o corregido.

1. Creación de variables (SPSS)

Utilizando la variable MES como base, recodificar dicha variable en una nueva denominada: MES T

Codificar los entes de control que reportan os casos a la ANT de 1 a 15, esto crea la variable ENTCONTROL. Seguido, asignar etiqueta a la variable.

Recodificar la variable original PROVINCIA para construir la nueva variable: PROVINCIA_T. Esto basado en el Clasificador Geográfico Estadístico – DPA. Seguido, asignar etiqueta a la variable

Se recodifica la variable DIA para generar la nueva variable DIA_T. Seguido, asignar etiqueta a la variable

Usando la variable HORA, se genera la variable HORA_T codificando las horas en rangos: 0 00:00 A 00:59 hasta 23 23:00 A 23:59. Seguido, asignar etiqueta a la variable

Generar la variable FERIADO_T donde se recodifique la variable original FERIADO donde: (SI=1) (NO=2). Seguido, asignar etiqueta a la variable

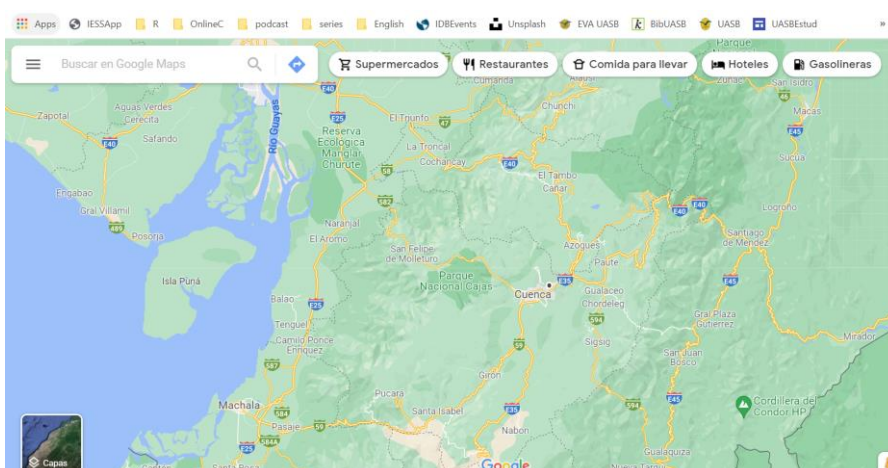
CAUSAPROBABLE: Esta variable posee varias causas probables que reporta la ANT. Se deben categorizar dentro de las causas que presenta INEC de forma homologada. La nueva variable es: CAUSA_T y presenta 10 categorías de las cerca de 27 causas que presenta ANT.

La variable CLASEFINAL tiene diferentes causas que se agrupan en las categorías que se presentan en INEC que son 8 categorías. De ese modo se crea la variable CLASE_T con las siguientes categorías:

- 1 ATROPELLOS
- 2 CAÍDA DE PASAJEROS
- 3 CHOQUES
- 4 ESTRELLAMIENTOS
- 5 ROZAMIENTOS
- 6 VOLCAMIENTOS
- 7 PÉRDIDA DE PISTA
- 8 OTROS

Recodificar la variable original ZONA en ZONA_T (URBANA=1) (RURAL=2). Seguido, asignar etiqueta a la variable

La variable LATITUDY es una variable de validación. Sí por alguna circunstancia se necesita determinar la ubicación exacta del siniestro, al igual que LONGITUDX, se pueden usar estas coordenadas en Google Maps para determinar el lugar exacto y realizar la corrección respectiva.



La variable CANTÓN se recodifica para construir la nueva variable: CANTON_T. Esto basado en el Clasificador Geográfico Estadístico – DPA.

Crear la variable: NUM_FALLECIDO_T que es igual a NUM_FALLECIDO.

Crear la variable: NUM_LESIONADO_T que es igual a NUM_LESIONADO. Para crear la variable: TOTAL_VICTIMAS, esta se genera como la suma de NUM_FALLECIDO_T y NUM_LESIONADO_T.

Posterior a la creación de variables se procede asignar etiquetas a las variables.

03.

Transporte Ferroviario

A pesar del cierre de la empresa se conserva la sintaxis de validación por si se reactiva en el futuro.

Proceso y metodología de las distintas etapas de validación de Transporte Ferroviario

1. Cambios y creación de variables

Usando como base la variable NombreFilial, categorizar los nombres de las filiales por las categorías:

- FILIAL 1 Centro
- FILIAL 2 Litoral
- FILIAL 3 Norte
- FILIAL 4 Sur.

La variable NombreViaje sirve para calcular el mes de la salida en ferrocarril. Generar una nueva variable denominada MES_SALIDA.

Seguido, crear la variable FECHA una vez se ha revisado las variables EDAD y FECHA DE NACIMIENTO.

Nota: En este punto se debe trabajar de cerca con la variable **FechaNacimiento** para calcular la edad exacta de la persona al momento del viaje.

Usando la variable: **Tipoidentificación** se puede corregir posibles errores al momento de calcular la edad, ya que suele pasar que existen problemas de digitación.

Así también, se puede usar la variable NombreCompleto, sólo para los pasajeros nacionales, ya que en otros casos no es posible dar con la fecha correcta.

Se extrae los datos de mes y día de la variable NombreViaje y se concatena con el año de referencia, ejem 01/02/2020

Para la variable País, se requiere recodificar la variable y homologar nombre de países.

Generar una nueva variable denominada SEXO, con las categorías:

- 1 Hombre
- 2 Mujer
- 9 No registrado.

04

Transporte Aéreo

Proceso y metodología de las distintas etapas de validación de Transporte Aéreo

Para este componente no existe una base de datos como tal. En su lugar se recibe una serie de tablas de transporte aéreo internacional y doméstico para pasajeros y carga. Con esta información se trabaja en replicar los cuadros de Excel que realiza la Dirección General de Aviación Civil.

CIUDADES	TOTAL	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
TOTAL	1.461.791	358.635	360.365	200.751	0	0	16.609	38.673	51.760	60.408	93.952	123.589	157.049
AMSTERDAN	48.893	16.796	16.853	10.121	0	0	1.964	1.973	3.934	4.951	4.898	4.218	6.185
ATLANTA	28.290	10.845	9.848	5.233	0	0	0	0	0	0	0	2.324	0
BOGOTA	162.034	47.558	51.758	27.578	0	0	0	258	0	0	7.536	10.414	16.932
BRASILIA	65	0	0	0	0	0	0	0	0	0	0	0	0
CALI	12.899	4.321	4.966	2.772	0	0	0	0	0	0	0	0	0
CANCUN	12.632	4.828	4.949	2.855	0	0	0	0	0	0	0	0	0
CARACAS	1.176	0	621	355	0	0	0	0	0	0	0	0	0
DALLAS	28.399	10.881	11.357	6.061	0	0	0	0	0	0	0	0	0
EL SALVADOR	24.176	7.828	10.825	5.523	0	0	0	0	0	0	0	0	0
FORT LAUDERDALE	139.594	25.629	23.733	14.012	0	0	4.328	8.066	9.046	9.788	12.129	13.810	19.053
HOUSTON	42.897	7.383	6.339	3.426	0	0	466	1.759	2.564	3.888	4.774	5.442	6.936
LA PAZ	293	0	136	86	0	0	0	0	0	0	0	0	0
LIMA	38.154	35.719	35.750	15.583	0	0	0	0	0	3.171	1.489	6.442	0
MADRID	143.993	38.441	31.899	18.529	0	0	645	9.178	9.183	8.592	7.825	7.987	11.714
MEXICO	66.135	16.552	18.230	9.220	0	0	0	2.991	4.135	4.597	4.887	5.523	0
MIAMI	220.914	32.645	33.107	18.837	0	0	8.972	12.039	16.788	18.446	21.965	28.725	0
NEW YORK	100.241	20.632	20.222	10.356	0	0	956	5.880	6.637	6.727	7.934	9.390	11.638
PANAMA	238.257	52.519	56.144	34.259	0	0	657	4.861	18.307	33.123	38.387	0	0
PARIS	12.439	4.867	4.463	3.109	0	0	0	0	0	0	0	0	0
RIO DE JANEIRO	60	0	58	0	0	0	0	0	0	0	0	0	0
SANTIAGO	39.813	11.767	11.541	8.002	0	0	178	0	0	816	4.319	3.190	0
SAO PAULO	7.248	3.282	2.545	1.421	0	0	0	0	0	0	0	0	0
TORONTO	14.779	5.918	5.416	3.445	0	0	0	0	0	0	0	0	0

Una vez que detectan los datos que se encuentran erróneos, se realiza una tabla donde se evidencie el error u omisión para enviar a la DGAC para su corrección o aclaración respectiva.

De lo contrario, se procede a generar los tabulados de este componente, para ello, se sugiere realizar un panel de los tabulados:

Tipo	Tipo mov	AEROP	ENTR/SAL	CIUDADES	TOTAL	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP
Pasajeros	Regular	GUAYAQU	ENTRADOS	AMSTERDAN	23.145	1.079	906	1.132	1.316	1.169	1.957	2.212	2.358	2.308
Pasajeros	Regular	GUAYAQU	ENTRADOS	BOGOTA	59.891	2.855	2.682	2.813	2.833	2.319	3.980	4.524	4.882	5.341
Pasajeros	Regular	GUAYAQU	ENTRADOS	CALI	73	0	0	0	0	0	0	0	0	0
Pasajeros	Regular	GUAYAQU	ENTRADOS	EL SALVADO	5.149	0	0	0	0	306	476	584	618	421
Pasajeros	Regular	GUAYAQU	ENTRADOS	FORT LAUDEI	79.683	3.998	3.679	2.828	3.753	7.013	8.202	9.362	7.368	7.451
Pasajeros	Regular	GUAYAQU	ENTRADOS	LIMA	16.980	1.333	855	850	797	1.066	993	1.185	1.529	1.314
Pasajeros	Regular	GUAYAQU	ENTRADOS	MADRID	50.415	1.455	1.363	1.356	776	2.151	4.635	6.692	6.003	6.060
Pasajeros	Regular	GUAYAQU	ENTRADOS	MIAMI	149.578	7.904	8.096	9.735	10.793	12.356	15.095	14.252	13.221	13.822
Pasajeros	Regular	GUAYAQU	ENTRADOS	NEW YORK	84.909	3.993	4.661	4.329	3.834	5.538	8.293	9.266	12.116	8.171
Pasajeros	Regular	GUAYAQU	ENTRADOS	ORLANDO	363	0	0	0	0	0	0	0	0	363
Pasajeros	Regular	GUAYAQU	ENTRADOS	PANAMA	197.177	7.771	8.504	9.466	10.104	14.439	15.723	20.012	21.252	20.698
Pasajeros	Regular	GUAYAQU	ENTRADOS	SANTIAGO	10.339	1.405	1.236	1.356	418	128	258	473	757	595
Pasajeros	Regular	GUAYAQU	SALIDOS	AMSTERDAN	25.998	1.547	315	1.075	1.640	1.233	1.424	1.946	2.856	3.943
Pasajeros	Regular	GUAYAQU	SALIDOS	BOGOTA	59.580	3.137	2.871	3.317	3.334	2.514	3.172	4.437	6.150	6.074

Al realizar éstas tablas unificando, por ejemplo, pasajeros entrados por cada aeropuerto de tráfico internacional, se facilita la validación y revisión de inconsistencias de sumas.

Además, facilita la combinación con tráfico doméstico para la construcción de la base para power bi.

Una vez realizado y corregido, de ser necesario, por la fuente, se puede generar tabulados.

05.

Transporte Marítimo

Proceso y metodología de las distintas etapas de validación de Transporte Marítimo

Las bases de transporte marítimo se reciben en formato Excel. Cada uno debe ser sometida a una revisión por separado, posteriormente se realiza la consolidación de las bases para ver todo en su conjunto y realizar validaciones.

La homologación de la información recibida por los diferentes puertos se debe realizar de acuerdo a las variables que disponga en las bases ya que algunos puertos, como las superintendencias petroleras, simplemente omiten o borran algunas variables. O, por otro lado, como APE incrementan variables. Una vez homologados, sobre todo en número de variables, se traslada a software con el que se trabajará.

Nota: para el año 2021 se incrementa las bases de puertos privados, 7 bases que conforman los Terminales Portuarios Habilitados – TPH.

1. Creación y validación de variables

Una vez se realiza la unión de bases en SPSS se puede iniciar la validación. No se realizan cambio y/o arreglos en esa unión de bases, simplemente se colocan de acuerdo al formato establecido ya que muchos puertos sólo envían las variables que ellos generan información.

NOMBREDELUERTO Homologar los nombres de los puertos:

- APE AUTORIDAD PORTUARIA DE ESMERALDAS.
- APG AUTORIDAD PORTUARIA DE GUAYAQUIL.
- APM AUTORIDAD PORTUARIA DE MANTA.
- APPB AUTORIDAD PORTUARIA DE PUERTO BOLÍVAR.
- SUINBA SUPERINTENDENCIA DEL TERMINAL PETROLERO DE BALAO.
- SUINLI SUPERINTENDENCIA DEL TERMINAL PETROLERO DE LA LIBERTAD.
- SUINSA SUPERINTENDENCIA DEL TERMINAL PETROLERO DE EL SALITRAL.
- **TPH TERMINALES PORTUARIOS HABILITADOS.**

TIPODEPUERTO: Variable de validación

IDREGISTRO: Variable de validación que sirve para completar datos del buque en caso de que falte alguna información. Se lo realiza en la página: <https://www.marinetraffic.com/>

IDBUQUE: Variable de validación que sirve para completar datos del buque en caso de que falte alguna información. Se lo realiza en la página: <https://www.marinetraffic.com/>

NRO.MATRÍCULA: Variable de validación que sirve para completar datos del buque en caso de que falte alguna información. Se lo realiza en la página: <https://www.marinetraffic.com/>

NOMBREDELANAVE: Variable de validación que sirve para completar datos del buque en caso de que falte alguna información. Se lo realiza en la página: <https://www.marinetraffic.com/>

TIPODELANAVE: Variable de validación

TRÁFICO: Se homologa el tipo de puerto:

string Codtraf(A15).

CABOTAJE o NACIONAL) = NACIONAL.

INTERNAC o INTERNACIONAL o Internacional = INTERNACIONAL: Se crea a una segunda base sólo de tráfico internacional para trabajar a partir de esa. Eso no quiere decir que la anterior no sea trabajada, ya que, se solicita las novedades sobre toda la información nacional o internacional

FECHAENTRADA: Se extrae sólo la fecha formato: DD/MM/YYYY sin horas. Se genera una variable auxiliar para comparación de fechas, que verifique que la fecha de entrada no sea mayor a la de salida

FECHASALIDA: Se extrae sólo la fecha formato: DD/MM/YYYY sin horas. Se genera una variable auxiliar para comparación de fechas, que verifique que la fecha de entrada no sea mayor a la de salida

BANDERA: Homologación de nombres de países. Se genera nueva variable Bandera con nombres de países homologados.

TONELAJEDEREGISTROBRUTOTRB: No se realizan cambios, sin embargo se revisa la información.

TONELAJEDEREGISTRONETOTRN: No se realizan cambios, sin embargo se revisa la información.

NÚMEROPASAJEROSENTRADOS: Se corroborar la información con lo remitido por Puertos. Generalmente APG y APM, son puertos de pasajeros, especialmente de cruceros, así que los pasajeros entrados deben ser los pasajeros que salieron. Se homologa pasajeros entrados y salidos.

NÚMEROPASAJEROSALIDOS: Se corroborar la información con lo remitido por Puertos. Generalmente APG y APM, son puertos de pasajeros, especialmente de cruceros, así que los pasajeros entrados deben ser los pasajeros que salieron. Se homologa pasajeros entrados y salidos.

AGENCIANAVIERA Se homologa los nombres de las agencias navieras. Surge dificultad ya que una sola empresa suele manejar varias agencias en Ecuador.

NÚMERODECONTENEDORES20ENT Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

NÚMERODECONTENEDORES20ENT Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

CARGACONTENEDORES20TMENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

NÚMERODECONTENEDORESVACIOS40ENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

NÚMERODECONTENEDORESLLENOS40ENT: Variable de publicación - se debe confirmar con suma total: Se crea variable auxiliar de comparación de sumas y totales

CARGACONTENEDORES40TMENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

TOTALDECARGACONTENERIZADATMENT: Variable de publicación - se debe confirmar con suma total: Se crea variable auxiliar de comparación de sumas y totales.

CARGAGRANELSÓLIDATMENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGAGRANELLÍQUIDATMENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGAGENERALTMENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGADETRANSBORDOTMENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

NÚMERODEVEHÍCULOSENT Variable de publicación - se debe confirmar con suma total incorporando a carga general ya que un solo puerto realiza esta diferenciación. Se crea variable auxiliar de comparación de sumas y totales.

VEHÍCULOSTMENT: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

TOTALCARGATMENT Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

NÚMERODECONTENEDORESVACIOS20SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

NÚMERODECONTENEDORESLLENOS20SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

CARGACONTENEDORES20TMSAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

NÚMERODECONTENEDORESVACIOS40SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales

NÚMERODECONTENEDORESLLENOS40SAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGACONTENEDORES40TMSAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

TOTALDECARGACONTENERIZADATMSAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGAGRANELSÓLIDATMSAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGAGRANELLÍQUIDATMSAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGAGENERALTMSAL: Variable de publicación - se deben confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

CARGADETRANSBORDOTMSAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

NÚMERODEVEHÍCULOSSAL: Variable de publicación - se debe confirmar con suma total incorporando a carga general ya que un solo puerto realiza esta diferenciación.

VEHÍCULOSTMSAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

TOTALCARGATMSAL: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

TOTLACARGAEXPORIMPOR: Variable de publicación - se debe confirmar con suma total. Se crea variable auxiliar de comparación de sumas y totales.

Zona: Esta variable tiene información para APG, permite diferenciar de los puertos públicos de privados. (Contecon, Andipuerto y DP World).

Observaciones generales

APE presenta una variable denominada: vehículos en Tm, variable que no estuvo en la matriz del año pasado enviada por el puerto, tanto en entrada como salida.

De acuerdo a la técnica del puerto, es para mayor detalle. Sin embargo, no se usa ese dato en validaciones.

APG presenta información de todos los puertos públicos, privados y bajo concesión que posee, lo que impide una depuración de la data al no poder identificar cuales conforman la data regular. Se espera las cifras del boletín de transporte marítimo para comparar y depurar de forma correcta.

5. Comparación de la información

Es importante comparar la información de cada componente con la fuente que los crea. Esto permite identificar errores y someterlos a un profundo análisis de los datos. Para esto, se contacta a las fuentes con las novedades presentadas y se solicita su revisión.

Para aquellos que disponen datos ya elaborados y publicados se realiza una comparación con estos:

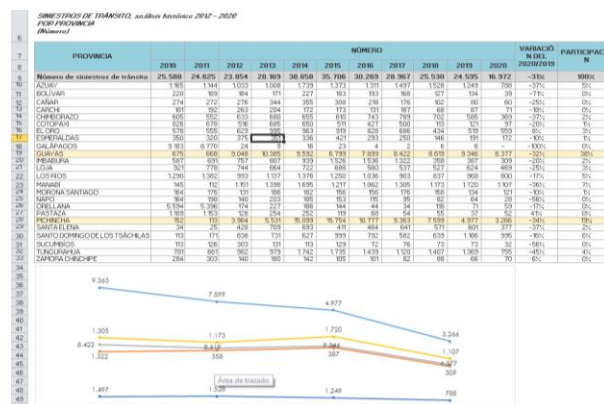
Ilustración 1 Documentos comparativos para las estadísticas de transporte



Fuente: Varios

Así también es necesario revisar la consistencia de los datos a través del tiempo.

Ilustración 2 Consistencia de las estadísticas histórica



Fuente: GESE

6.

- Este documento es de ayuda, ya que, detalla las diferentes etapas de validación a los que es sometida la información recibida desde las diferentes fuentes del Anuario y, dicta las pautas para procedimientos futuros así como la mejora de estos.
- Es importante generar procesos de validación para prevenir errores en la información recibida. Es así que, la planificación de la recolección es clave para tener datos oportunos y que estos puedan ser sometidos a las validaciones para identificar errores y tener retroalimentaciones con las fuentes de los registros administrativos para obtener datos de calidad.
- El trabajo conjunto entre el equipo técnico y los técnicos de las diferentes instituciones fuente, ha favorecido al momento de identificar y explicar datos que pueden estar inconsistentes. De esta manera, se obtienen cifras homologadas entre las instituciones que, sobre todo ayudan al usuario.
- Después de realizadas las validaciones, se generan resultados oportunos y de calidad, así como otros productos para la revisión y probación de las diferentes instancias del INEC. Finalmente, estos son presentados a los usuarios del anuario de acuerdo al calendario estadístico.

Actualizado al: 5 de oct. de 22.

Elaborado por:	Revisado y Aprobado por:
Gabriela Hidalgo Miembro del Equipo	Roberto Chaves Jefe de área

